JOINT FUNDAMENTAL FREQUENCY AND ORDER ESTIMATION USING OPTIMAL FILTERING

Mads G. Christensen[1] , *Jesper Højvang Jensen*[1], *Andreas Jakobsson*[2], *and Søren Holdt Jensen*[1]

[1] Dept. of Electronic Systems, Aalborg University, Denmark
{mgc,jhj,shj}@es.aau.dk
[2] Dept. of Mathematical Statistics, Lund University, Sweden
andreas.jakobsson@ieee.org

ABSTRACT

Recently, two optimal filter designs for fundamental frequency estimation have been proposed with the first being based on a filterbank and the second on a single filter. The two designs are related in a simple manner and are shown to result in the same residual when used for cancelling out the harmonics of periodic signals. We propose to use this residual for estimating the number of harmonics by combining a noise variance estimate with an order dependent penalty term. This leads to a joint estimator of the fundamental frequency and the order based on the same criterion. Via Monte Carlo simulations, the estimator is demonstrated to have good performance in terms of the percentage of correctly estimated orders.

1. INTRODUCTION

Periodic signals consist of a set of sinusoids whose frequencies are integer multiples of a fundamental frequency. The task of finding this fundamental frequency from an observed signal is important in applications for many kinds of signals, but especially so for speech and audio signals. Fundamental frequency estimators form the basis of many signal processing applications with some examples being separation, compression, analysis, and enhancement. The different methods for fundamental frequency estimation are too numerous to mention in any detail here and we will refer to [1] for an overview of classical methods. Some examples of more recent methods based on estimation theoretical approaches are [2–4]. One particular method, though, is the main inspiration for this paper, namely the adaptive comb filtering method [5]. We will here pursue the idea of obtaining fundamental frequency estimates using filters further, but unlike the fixed filter design of [5], we will use the signal-adaptive and optimal filters proposed in [6] and [7]. More specifically, two optimal filter designs for finding the fundamental frequency from an observed signal were proposed. The first design is based on a filterbank and was first introduced in [6] where it was shown to have excellent performance compared to many other methods, while the second design is new and is based on a single filter [7]. Both are generalizations of Capon's classical filter design [8] that has been used extensively for spectral estimation and beamforming in array processing. In this paper, we extend these estimators to also account for an unknown number of harmonics, something that is critical in avoiding ambiguities in the cost function (see, e.g., [9] for more on this) by deriving a method for estimating the fundamental frequency and the number of harmonics and the model order jointly for a particular data segment. In doing so, a noise variance estimate is obtained, and we show that the resulting optimal filterbank and the single filter lead to identical noise variance estimates.

M. G. Christensen is supported by the Parametric Audio Processing project, Danish Research Council for Technology and Production Sciences grant no. 274–06–0521.

We will make use of the following signal model and notation: a signal consisting of a set of sinusoids having frequencies that are integer multiples of a fundamental frequency, ω_0, is corrupted by an additive white complex circularly symmetric Gaussian noise, $e(n)$, having variance σ^2, for $n = 0,\dots,N-1$, i.e.,

$$x(n) = \sum_{l=1}^{L} \alpha_l e^{j\omega_0 l n} + e(n), \tag{1}$$

where $\alpha_l = A_l e^{j\psi_l}$, with $A_l > 0$ and ψ_l being the amplitude and the phase of the lth harmonic, respectively. The problem of interest is to estimate the fundamental frequency ω_0 as well as the order L from a set of N measured samples $x(n)$.

The remaining part of the present paper is organized as follows: First, we introduce two different filter designs in Section 2, one based on a filterbank and one based on a single filter. Then, in Section 3, we will derive a noise variance estimator for estimating the order based on these filter designs. In Section 4, we then proceed to evaluate the proposed joint fundamental frequency and order estimator before concluding on our work in Section 5.

2. OPTIMAL DESIGNS

2.1 Filterbank

We begin by introducing some useful notation and definitions. First, we introduce a vector formed from M time-reversed samples of the observed signal, i.e., $\mathbf{x}(n) = [\, x(n)\ x(n-1)\ \cdots\ x(n-M+1)\,]^T$ with $M \le N/2$ and with $(\cdot)^T$ denoting the transpose. Next, we define the output signal $y_l(n)$ of the lth filter having coefficients $h_l(n)$ as

$$y_l(n) = \sum_{m=0}^{M-1} h_l(m) x(n-m) = \mathbf{h}_l^H \mathbf{x}(n), \tag{2}$$

with $\mathbf{h}_l = [\, h_l(0)\ \cdots\ h_l(M-1)\,]^H$ and $(\cdot)^H$ denoting the Hermitian transpose. Introducing the expected value $\mathrm{E}\{\cdot\}$ and defining the covariance matrix as $\mathbf{R} = \mathrm{E}\{\mathbf{x}(n)\mathbf{x}^H(n)\}$, the output power of the lth filter of the filterbank can be expressed as

$$\mathrm{E}\left\{|y_l(n)|^2\right\} = \mathrm{E}\left\{\mathbf{h}_l^H \mathbf{x}(n)\mathbf{x}^H(n)\mathbf{h}_l\right\} = \mathbf{h}_l^H \mathbf{R} \mathbf{h}_l. \tag{3}$$

The sum of the output powers of all the filters is given by

$$\sum_{l=1}^{L} \mathrm{E}\left\{|y_l(n)|^2\right\} = \sum_{l=1}^{L} \mathbf{h}_l^H \mathbf{R} \mathbf{h}_l = \mathrm{Tr}\left[\mathbf{H}^H \mathbf{R} \mathbf{H}\right], \quad (4)$$

where $\mathbf{H}$ contains the filters, i.e., $\mathbf{H} = [\, \mathbf{h}_1 \cdots \mathbf{h}_L \,]$. An intuitive approach is to seek to find a set of filters that pass power undistorted at specific frequencies, in our case the harmonic frequencies, while minimizing the power at all other frequencies. This problem can be formulated mathematically as the optimization problem:

$$\min_{\mathbf{H}} \mathrm{Tr}\left[\mathbf{H}^H \mathbf{R} \mathbf{H}\right] \quad \text{s.t.} \quad \mathbf{H}^H \mathbf{Z} = \mathbf{I}, \quad (5)$$

where $\mathbf{I}$ is the $L \times L$ identity matrix. Furthermore, the matrix $\mathbf{Z} \in \mathbb{C}^{M \times L}$ has a Vandermonde structure and is constructed from L complex sinusoidal vectors as

$$\mathbf{Z} = [\, \mathbf{z}(\omega_0) \cdots \mathbf{z}(\omega_0 L) \,], \quad (6)$$

with $\mathbf{z}(\omega) = [\, 1 \; e^{-j\omega} \cdots e^{-j\omega(M-1)} \,]^T$, i.e., the matrix contains the harmonically related complex sinusoids. We note that the complex conjugation is due to the covariance matrix being defined from the time-reversed signal vector $\mathbf{x}(n)$. Using the method of Lagrange multipliers, the unconstrained optimization problem can be written as

$$\mathscr{L}(\{\mathbf{h}_l\}, \{\lambda_l\}) = \sum_{l=1}^{L} \mathbf{h}_l^H \mathbf{R} \mathbf{h}_l - \left(\mathbf{h}_l^H \mathbf{Z} - \mathbf{b}_l^T\right) \lambda_l \quad (7)$$

with $[\mathbf{b}_l]_v = 0$ for $v \neq l$ and $[\mathbf{b}_l]_v = 1$ for $v = l$, i.e., each individual filter is constrained to have unit gain for a certain harmonic frequency and zero gain for the others. It is easy to see that this can be written using a more convenient form as

$$\mathscr{L}(\mathbf{H}, \Lambda) = \mathrm{Tr}\left\{\mathbf{H}^H \mathbf{R} \mathbf{H}\right\} - \mathrm{Tr}\left\{\left(\mathbf{H}^H \mathbf{Z} - \mathbf{I}\right) \Lambda\right\}, \quad (8)$$

where the matrix Λ contains all the Lagrange multiplier (column) vectors λ_l associated with the various filters of the filterbank, i.e.,

$$\Lambda = [\, \lambda_1 \cdots \lambda_L \,]. \quad (9)$$

By differentiation, we obtain that the gradient of this composite cost function is

$$\nabla \mathscr{L}(\mathbf{H}, \Lambda) = \begin{bmatrix} \mathbf{R} & -\mathbf{Z} \\ -\mathbf{Z}^H & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{H} \\ \Lambda \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}. \quad (10)$$

By setting these matrix equations equal to zero, one readily obtains that the Lagrange multipliers that solve the original problem are

$$\Lambda = \left(\mathbf{Z}^H \mathbf{R}^{-1} \mathbf{Z}\right)^{-1} \quad (11)$$

and that the optimal filterbank expressed in terms of the Lagrange multipliers is

$$\mathbf{H} = \mathbf{R}^{-1} \mathbf{Z} \Lambda. \quad (12)$$

By substituting the solution for the Lagrange multipliers, the filter bank matrix $\mathbf{H}$ solving (5) can be seen to be given by

$$\mathbf{H} = \mathbf{R}^{-1} \mathbf{Z} \left(\mathbf{Z}^H \mathbf{R}^{-1} \mathbf{Z}\right)^{-1}. \quad (13)$$

This data and frequency dependent filter bank can then be used to estimate the fundamental frequencies by treating it as an unknown variable and maximizing the power of the filter's output, which is

$$\hat{\omega}_0 = \arg\max_{\omega_0} \mathrm{Tr}\left[\left(\mathbf{Z}^H \mathbf{R}^{-1} \mathbf{Z}\right)^{-1}\right]. \quad (14)$$

This expression depends only on the covariance matrix and the Vandermonde matrix constructed for different candidate fundamental frequencies.

2.2 Single Filter

There is an alternative formulation of the filter design problem that we will now examine further. Suppose that we instead wish to design a single filter $\mathbf{h}$ that passes the signal undistorted at the harmonic frequencies and suppresses everything else. This filter design problem can be stated mathematically as

$$\min_{\mathbf{h}} \mathbf{h}^H \mathbf{R} \mathbf{h} \quad \text{s.t.} \quad \mathbf{h}^H \mathbf{z}(\omega_0 l) = 1, \quad \text{for} \quad l = 1, \ldots, L. \quad (15)$$

It is worth noting that the single filter in (15) is designed subject to L constraints, whereas in (5) the filter bank is designed using a number of constraints for each filter. Clearly, these two formulations are related; we will return to this relation later on. First, we will derive the optimal filter. Introducing the Lagrange multiplier column vector λ, the Lagrangian dual function associated with the problem stated above can be written as

$$\mathscr{L}(\mathbf{h}, \lambda) = \mathbf{h}^H \mathbf{R} \mathbf{h} - \left(\mathbf{h}^H \mathbf{Z} - \mathbf{1}^T\right) \lambda \quad (16)$$

with $\mathbf{1} = [\, 1 \cdots 1 \,]^T$. Taking the derivative with respect to the unknown filter impulse response, $\mathbf{h}$, as well as the Lagrange multipliers, we get

$$\nabla \mathscr{L}(\mathbf{h}, \lambda) = \begin{bmatrix} \mathbf{R} & -\mathbf{Z} \\ -\mathbf{Z}^H & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{h} \\ \lambda \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}. \quad (17)$$

By setting this expression equal to zero, i.e., $\nabla \mathscr{L}(\mathbf{h}, \lambda) = \mathbf{0}$, and solving for the unknowns, we obtain, as with the filterbank design, the optimal Lagrange multipliers for which the equality constraints are satisfied as

$$\lambda = \left(\mathbf{Z}^H \mathbf{R}^{-1} \mathbf{Z}\right)^{-1} \mathbf{1} \quad (18)$$

and the optimal filter as

$$\mathbf{h} = \mathbf{R}^{-1} \mathbf{Z} \lambda. \quad (19)$$

By combining the last two expressions, we get the optimal filter expressed in terms of the covariance matrix and the Vandermonde matrix $\mathbf{Z}$, i.e.,

$$\mathbf{h} = \mathbf{R}^{-1} \mathbf{Z} \left(\mathbf{Z}^H \mathbf{R}^{-1} \mathbf{Z}\right)^{-1} \mathbf{1}. \quad (20)$$

The output power of this filter can then be expressed as

$$\mathbf{h}^H \mathbf{R} \mathbf{h} = \mathbf{1}^H \left(\mathbf{Z}^H \mathbf{R}^{-1} \mathbf{Z}\right)^{-1} \mathbf{1}, \quad (21)$$

which, as for the first design, depends only on the inverse of $\mathbf{R}$ and the Vandermonde matrix $\mathbf{Z}$. By maximizing the output power, we readily obtain an estimate of the fundamental frequency as

$$\hat{\omega}_0 = \arg\max_{\omega_0} \mathbf{1}^H \left(\mathbf{Z}^H \mathbf{R}^{-1} \mathbf{Z}\right)^{-1} \mathbf{1}. \quad (22)$$

2.3 Properties

The question arises as to exactly how the two approaches differ. Comparing the optimal filters in (13) and (20), it can be observed that the latter can be written in terms of the former as

$$\mathbf{h} = \mathbf{R}^{-1}\mathbf{Z}\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}\mathbf{1} = \mathbf{H}\mathbf{1} = \sum_{l=1}^{L}\mathbf{h}_l, \qquad (23)$$

so, clearly, the two methods are related, but on the other hand

$$\mathbf{1}^H\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}\mathbf{1} \neq \mathrm{Tr}\left[\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}\right], \qquad (24)$$

with equality only when $\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}$ is diagonal. For white noise where $\mathbf{R}^{-1} = \frac{1}{\sigma^2}\mathbf{I}$, the optimal filterbank and single filter reduce to

$$\mathbf{H} = \mathbf{Z}\left(\mathbf{Z}^H\mathbf{Z}\right)^{-1} \quad \text{and} \quad \mathbf{h} = \mathbf{Z}\left(\mathbf{Z}^H\mathbf{Z}\right)^{-1}\mathbf{1}, \qquad (25)$$

respectively. In Figure 1, an example of such filters are given with the magnitude response of the optimal filterbank and the single filter being shown for white Gaussian noise with $\omega_0 = 1.2566$ and $L = 3$. It should be stressed that for a non-diagonal $\mathbf{R}$, the resulting filters will look radically different. Interestingly, the two methods can be shown to be asymptotically equivalent as $\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}$ is in fact diagonal asymptotically in M when normalized appropriately in the sense that [7]

$$\lim_{M\to\infty} M\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1} = \begin{bmatrix} \Phi(\omega_0) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \Phi(\omega_0 L) \end{bmatrix},$$

where $\Phi(\omega)$ is the power spectral density of the observed signal $x(n)$, which has been assumed to be non-zero and finite[1]. This means that the two cost functions are generally different for small N and M and may result in different fundamental frequency estimates, but asymptotically they tend to the same cost function.

For the optimal filtering methods, the choise of the filter length M requires some consideration. It can be seen that both methods require that $\mathbf{R}$ be invertible, which is always the case for the signal model considered here when $\mathbf{R}$ is defined by the expectation operator. However, when the sample covariance matrix is used in its place, M must be chosen such that the sample covariance matrix has rank M, i.e., $M \leq N/2$. To obtain a good estimate of the covariance matrix and thus an accurate output power estimate, M should be chosen low so that many sub-vectors are used in the averaging. On the other hand, a high M results in more selective filters.

3. ORDER ESTIMATION

In order to determine the order, i.e., the number of harmonics, we employ the maximum a posteriori (MAP) principle [10, 11]. The derivations of the MAP criterion are somewhat lengthy and we will therefore only present the results here.

[1] This is strictly speaking not the case for the signal model in (1). Nevertheless, the results obtained under these assumptions still provide some useful insights into the properties of the methods.

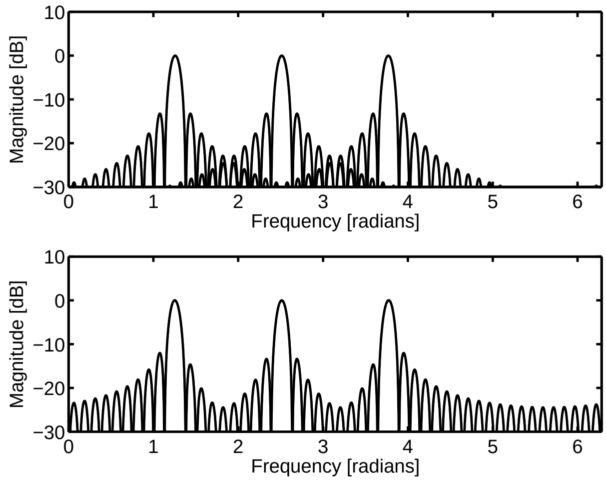

Figure 1: Magnitude response of optimal filterbank (top) and single filter (bottom) for white noise with $\omega_0 = 1.2566$ and $L = 3$.

More specifically, the MAP criterion for determining $L \geq 1$ in (1) can be shown to be

$$\hat{L} = \arg\min_{L} N \log \hat{\sigma}_L^2 + L \log N + \frac{3}{2}\log N, \qquad (26)$$

where the first term is a log-likelihood term that comprised a noise variance estimate that depends on the candidate model order L, the second is the penalty associated with the amplitude and phase of (1) while the third term is due to the fundamental frequency. Note that the linear and nonlinear parameters have different penalties associated with them. To determine whether any harmonics are present at all, the above cost function should be compared to the log-likelihood of the zero order model, meaning that no harmonics are present if

$$N \log \hat{\sigma}_0^2 < N \log \hat{\sigma}_{\hat{L}}^2 + \hat{L}\log N + \frac{3}{2}\log N. \qquad (27)$$

We will now proceed to use the filters presented in Section 2 to estimate the variance of the signal once the harmonics have been filtered out. First, we will do this based on the filterbank design. An estimate of the noise is defined as $\hat{e}(n) = x(n) - y(n)$ which we will refer to as the residual. Additionally, $y(n)$ is the sum of the input signal filtered by the filterbank, i.e.,

$$y(n) = \sum_{m=0}^{M-1}\sum_{l=1}^{L} h_l(m)x(n-m) = \sum_{m=0}^{M-1} h(m)x(n-m), \qquad (28)$$

where $h(m)$ is the sum over the impulse response of the filters of the filterbank. From the relation between the single filter design and the filterbank design in (23), it is now clear that when used this way, the two approaches lead to the same output signal $y(n)$. This also offers some insights into the difference between the designs (5) and (15). More specifically, the difference is in terms of the way the output power is measured, where (5) is based on the assumption that the power is additive over the filters. We can now write the noise estimate as

$$\hat{e}(n) = x(n) - \sum_{m=0}^{M-1} h(m)x(n-m) \triangleq \mathbf{g}^H\mathbf{x}(n) \qquad (29)$$

where $\mathbf{g} = [\ (1-h(0))\ \ -h(1)\ \cdots\ -h(M-1)\ ]^H$ is the modified filter. From the noise estimate, we can now estimate the noise variance for the Lth order model as

$$\hat{\sigma}_L^2 = \mathrm{E}\left\{|\hat{e}(n)|^2\right\} = \mathrm{E}\left\{\mathbf{g}^H\mathbf{x}(n)\mathbf{x}^H(n)\mathbf{g}\right\} = \mathbf{g}^H\mathbf{R}\mathbf{g}. \quad (30)$$

This expression is however not very convenient for a number of reasons. A notable property of the estimator in (22) is that it does not require the calculation of the filter and that the output power expression in (21) is simpler than the expression for the optimal filter in (20). To use (30) directly, we would first have to calculate the optimal filter using (20), then calculate the modified filter $\mathbf{g}$, before evaluating (30). Instead, we propose to simplify the evaluation of (30) by defining the modified filter as $\mathbf{g} = \mathbf{b}_1 - \mathbf{h}$ where, as defined earlier, $\mathbf{b}_1 = [\ 1\ 0 \cdots 0\ ]$. Next, we use this definition to rewrite the variance estimate as

$$\hat{\sigma}_L^2 = \mathbf{g}^H\mathbf{R}\mathbf{g} = (\mathbf{b}_1 - \mathbf{h})^H\mathbf{R}(\mathbf{b}_1 - \mathbf{h}) \quad (31)$$

$$= \mathbf{b}_1^H\mathbf{R}\mathbf{b}_1 - \mathbf{b}_1^H\mathbf{R}\mathbf{h} - \mathbf{h}^H\mathbf{R}\mathbf{b}_1 + \mathbf{h}^H\mathbf{R}\mathbf{h}. \quad (32)$$

The first term can be identified to equal the variance of the observed signal $x(n)$, i.e., $\mathbf{b}_1^H\mathbf{R}\mathbf{b}_1 = \mathrm{E}\left\{|x(n)|^2\right\}$ and $\mathbf{h}^H\mathbf{R}\mathbf{h}$ we know from (21). Writing out the cross-terms $\mathbf{b}_1^H\mathbf{R}\mathbf{h}$ using (20) yields

$$\mathbf{b}_1^H\mathbf{R}\mathbf{h} = \mathbf{b}_1^H\mathbf{R}\mathbf{R}^{-1}\mathbf{Z}\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}\mathbf{1} \quad (33)$$

$$= \mathbf{b}_1^H\mathbf{Z}\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}\mathbf{1}. \quad (34)$$

Furthermore, it can easily be verified that $\mathbf{b}_1^H\mathbf{Z} = \mathbf{1}^H$, from which it can be concluded that

$$\mathbf{b}_1^H\mathbf{R}\mathbf{h} = \mathbf{1}^H\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}\mathbf{1} = \mathbf{h}^H\mathbf{R}\mathbf{h}. \quad (35)$$

Therefore, the variance estimate can be expressed as

$$\hat{\sigma}_L^2 = \mathrm{E}\left\{|x(n)|^2\right\} - \mathbf{1}^H\left(\mathbf{Z}^H\mathbf{R}^{-1}\mathbf{Z}\right)^{-1}\mathbf{1}, \quad (36)$$

which conveniently features the same expression as in the fundamental frequency estimation criterion in (22). This means that the same expression can be used for determining the model order and the fundamental frequency, i.e., the approach allows for joint estimation of the model order and the fundamental frequency. It also shows that the same filter that maximizes the output power minimizes the variance of the residual. A more conventional variance estimate could be formed by first finding the frequency using, e.g., (22) and then finding the amplitudes of the signal model using least-squares to obtain a noise variance estimate. Since the proposed procedure uses the same information in finding the fundamental frequency and the noise variance, it is superior to the least-squares approach in terms of computational complexity. Note that for finite filter lengths, the output of the filters considered here are generally "power levels" and not power spectral densities (see [12]), which is consistent with our use of the filters. Asymptotically, the filters do comprise power spectral density estimates [7].

4. RESULTS

We will now evaluate the statistical performance of the proposed scheme. In doing so, we will compare to two other methods based on well-established estimation theoretical approaches that are able to jointly estimate the fundamental frequency and the order, namely a subspace method, the MUSIC method of [9], and the nonlinear least-squares (NLS) method [6]. The NLS method in combination with the criterion (26) yields both a maximum likelihood fundamental frequency estimate and a MAP order estimate (see [10] for details). The three methods are comparable in terms of computational efficiency as they all have complexity $\mathcal{O}(M^3)$. We will here focus on their application to order estimation, investigating the performance of the estimators given the fundamental frequency. The reason for this is simply that the high-resolution estimation capabilities of the proposed method, MUSIC and NLS for the fundamental frequency estimation problem are already well-documented in [6, 7, 9]. Note that the NLS method reduces to a linear least-squares method when the fundamental frequency is given but the joint estimator is still nonlinear. In these experiments the following conditions were used: signals were generated using (1) with a fundamental frequency of $\omega_0 = 0.8170$, $L = 5$ and $A_l = 1\ \forall l$. For each test condition, 1000 Monte Carlo iterations were run. In the first experiment, we will investigate the performance as a function of the pseudo signal-to-noise (PSNR) as defined in [9]. Note that this PSNR is higher than the usual SNR, meaning that the conditions are more noisy than they may appear at first sight. The performance of the estimators has been evaluated for $N = 200$ observed samples with a covariance matrix size/filter length of $M = 50$. The results are shown in Figure 2 in terms of the percentage of correctly estimated orders. Similarly, the performance is investigated as a function of N with $M = N/4$ in the second experiment for $PSNR = 40$ dB, i.e., the filter length is set proportionally to the number of samples. Note that the NLS method operates on the entire length N signal and thus does not depend on M. This experiment thus reveals not only the dependency of the performance on the number of observed samples but also on the filter length. The results are shown in Figure 3. In the final experiment, the N is kept fixed while the filter length M is varied with $PSNR = 40$ dB. In the process, the covariance matrix of MUSIC is varied too. The results can be seen in figure Figure 4. From the figures, it can be observed that

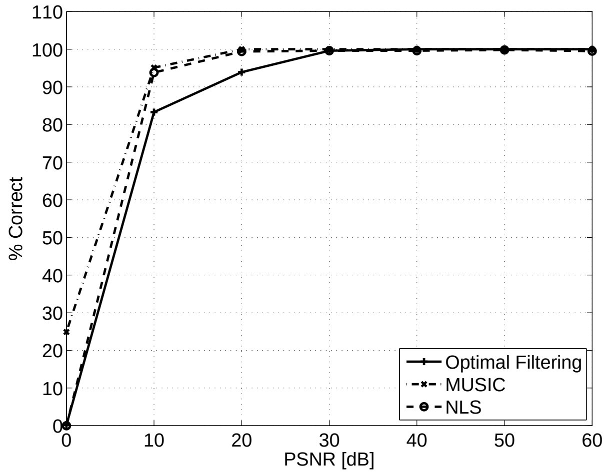

Figure 2: Percentage of correctly estimated model orders as a function of the PSNR.

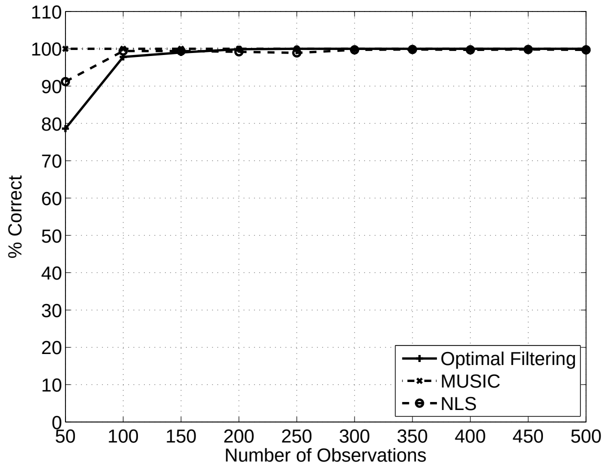

Figure 3: Percentage of correctly estimated model orders as a function of the number of samples N.

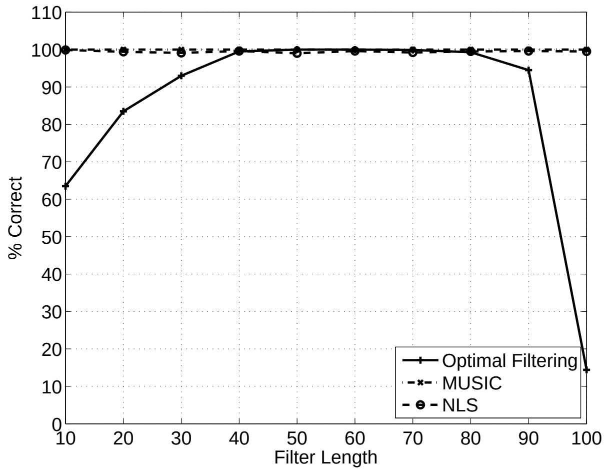

Figure 4: Percentage of correctly estimated model orders as a function of the filter length M.

the proposed method has good performance for high PSNRs and N with the percentage approaching 100 %. Furthermore, the filter length should not be chosen too low or too close to $N/2$. It can, however, also be observed that the method appears to be more sensitive than MUSIC and NLS to low number of samples (and thus filter lengths) and SNRs. It should be stressed, though, that while the method based on optimal filtering appears to generally exhibit slightly worse performance than both MUSIC and NLS in terms of estimating the model order, it generally outperforms both MUSIC and the NLS with respect to fundamental frequency estimation under adverse conditions, in particular when multiple periodic sources are present at the same time [6], something that happens frequently in audio signals. Also, it should be noted that for our intended application, which is audio processing, the segment lengths are generally high. More specifically, segments of 30 ms or longer sampled at 44.1 kHz corresponding to 1323 samples are commonly used in audio estimation problems. That the proposed method appears to require long filters is therefore not necessarily a concern.

5. CONCLUSION

Two optimal filter designs that can be used for estimating the fundamental frequency of a periodic signal have been considered. Both are based on Capon's classical filter design and are related in a simple way. In this paper, we have extended the principles of these methods to also account for an unknown model order leading to a joint estimator for the fundamental frequency and the number of harmonics. In Monte Carlo simulations, the proposed scheme is demonstrated to have good performance estimating the correct order with a high probability for a high number of observations. The results are promising as methods based on optimal filtering have previously been shown to have excellent performance even under adverse conditions for a known model order. The results are particularly relevant for speech and audio signals where the model order may vary greatly over time.

REFERENCES

[1] W. Hess, *Pitch Determination of Speech Signals*, Springer-Verlag, Berlin, 1983.

[2] S. Godsill and M. Davy, "Bayesian harmonic models for musical pitch estimation and analysis," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 2002, vol. 2, pp. 1769–1772.

[3] K. W. Chan and H. C. So, "Accurate frequency estimation for real harmonic sinusoids," *IEEE Signal Processing Lett.*, vol. 11(7), pp. 609–612, July 2004.

[4] H. Li, P. Stoica, and J. Li, "Computationally efficient parameter estimation for harmonic sinusoidal signals," *Signal Processing*, vol. 80, pp. 1937–1944, 2000.

[5] A. Nehorai and B. Porat, "Adaptive comb filtering for harmonic signal enhancement," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34(5), pp. 1124–1138, Oct. 1986.

[6] M. G. Christensen, P. Stoica, A. Jakobsson and S. H. Jensen, "Multi-pitch estimation," *Elsevier Signal Processing*, vol. 88(4), pp. 972–983, Apr. 2008.

[7] M. G. Christensen, J. H. Jensen, A. Jakobsson and S. H. Jensen, "On optimal filter designs for fundamental frequency estimation," *IEEE Signal Processing Lett.*, vol. 15, pp. 745–748, 2008.

[8] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57(8), pp. 1408–1418, 1969.

[9] M. G. Christensen, A. Jakobsson and S. H. Jensen, "Joint high-resolution fundamental frequency and order estimation," *IEEE Trans. on Audio, Speech and Language Processing*, vol. 15(5), pp. 1635–1644, July 2007.

[10] P. M. Djuric, "Asymptotic MAP criteria for model selection," *IEEE Trans. Signal Processing*, vol. 46, pp. 2726–2735, Oct. 1998.

[11] P. Stoica and Y. Selen, "Model-order selection: a review of information criterion rules," *IEEE SP Mag.*, vol. 21(4), pp. 36–47, July 2004.

[12] M. A. Lagunas, M. E. Santamaria, A. Gasull, and A. Moreno, "Maximum likelihood filters in spectral estimation problems," *Elsevier Signal Processing*, vol. 10(1), pp. 19–34, 1986.